# Measure and exploitation of multisensor and multiwavelength synergy for remote sensing: 1. Theoretical considerations

Filipe Aires

# Measure and exploitation of multisensor and multiwavelength synergy for remote sensing: 1. Theoretical considerations

Filipe Aires[1,2,3]

Received 30 June 2010; revised 22 October 2010; accepted 12 November 2010; published 20 January 2011.

[1] A synergetic scheme refers to an algorithm that simultaneously or hierarchically uses the observations of two or more spectral ranges in order to obtain a more accurate retrieval than the independent retrievals put together. This study is composed of two companion papers; this first part introduces some theoretical considerations. The goal of this study is, first, to identify the various forms of synergy for remote sensing applications. Simple linear models are used to introduce concepts such as additive, unmixing, indirect, or denoising synergies. The second objective of this paper is to develop a methodology to measure, in real-world applications, these different synergies. For this purpose, some experiments are conducted using the classical information content analysis which is often used in the context of assimilation or to design new instruments. This technique is tested on a real-world application where the microwave and infrared observations from the Atmospheric Microwave Sounding Unit-A, Microwave Humidity Sounder, and Improved Atmospheric Sounding in the Infrared instruments are used to retrieve the atmospheric profiles of temperature and water vapor over ocean, under clear-sky conditions. This approach will show its limitation to measure synergy and stress the need for other tools. In the companion paper, statistical retrieval schemes will show their potential to measure and exploit existing synergies, for the same application.

**Citation:** Aires, F. (2011), Measure and exploitation of multisensor and multiwavelength synergy for remote sensing: 1. Theoretical considerations, *J. Geophys. Res.*, *116*, D02301, doi:10.1029/2010JD014701.

## 1. Introduction

[2] A wealth of Earth satellite observations is now available, covering the entire globe, and providing a large diversity of information over a broad frequency range (UV, visible, infrared, microwave), in order to obtain a global and continuous monitoring of the state of the atmosphere. Space agencies have designed satellite platforms that include instruments from the different regions of the electromagnetic spectrum. In parallel, accurate Radiative Transfer Models (RTM) have been developed to simulate the responses of these multispectral observations to atmospheric changes in composition or temperature. However, the retrieval accuracy of key variables such as temperature or water vapor profiles is still not always satisfying.

[3] The radiation measured by an instrument on board a satellite often results from the combination of contributions from the different atmospheric constituents, including gases, aerosols, and hydrometeors, as well as possibly from the Earth surface. Disentangling the various effects in order to quantify a given atmospheric or surface variable can be very challenging. Historically, methodologies have been developed to use one type of wavelengths to derive one atmospheric parameter. To suppress ambiguities related to the contribution from other effects, these algorithms exploit for instance the complementarity of close frequency bands measured by the same instrument or use ancillary information from independent sources. Simultaneous observations of the atmospheric state in different wavelength ranges could help separate the different contributions in order to obtain better estimates of a given parameter. However, despite the availability of multifrequency satellite observations, limited efforts have been invested so far to design retrieval schemes that benefit from the potential synergy between the measurements from different wavelength ranges. In the work by *Cho and Staelin* [2006], Atmospheric Microwave Sounding Unit (AMSU) data are used to perform a cloud clearing of the AIRS infrared radiances. MODIS and AIRS measurements are exploited in a variational retrieval scheme for cloud parameters [*Li et al.*, 2004], and *Susskind et al.* [2003] combine AIRS/AMSU/HSB data for the retrieval of atmospheric and surface parameters in the presence of clouds. Most of the time, retrievals from different instruments are combined a posteriori, and the potential synergy between the observations is not exploited. A synergetic scheme refers to an algorithm that uses

[1]CNRS, IPSL, Laboratoire de Météorologie Dynamique, Université Pierre et Marie Curie, Paris, France.
[2]Also at Laboratoire de l'Etude du Rayonnement et de la Matière en Astrophysique, CNRS, Observatoire de Paris, Paris, France.
[3]Now at Estellus, Paris, France.


0148-0227/11/2010JD014701

simultaneously or hierarchically the observations of two or more spectral ranges in order to obtain a more accurate retrieval than the independent retrievals together. The definition of synergy is extended here to the inversion schemes that combine information (observations, a priori information or retrieved parameters) in order to improve their retrievals. We define the synergy factor of a retrieval scheme using $R$ sources of information $(x_1,\ldots,x_R)$, each one can be multivariate, as the ratio of the errors of the retrieval using the best single information, $\min_{i=1,\ldots,R} E(x_i)$ with the errors of the retrieval using all the sources of information, $E(x_1,\ldots,x_R)$:

$$F_{\text{syn}} = \frac{\min_{i=1,\ldots,R} E(x_i)}{E(x_1,\ldots,x_R)}.$$

There is synergy when this ratio is >100%. How are different sources of information used efficiently for a better final product? The retrieval scheme has to merge different satellite observations, to optimize the use of complementary observations, and to exploit potential a priori information. Numerical Weather Prediction (NWP) assimilation models are able to use multispectral observations. However, it is difficult with this approach to actually assess and evaluate the potential synergy and thereby identify measurement improvements. There is still place for improvement in the retrieval of key atmospheric variables such as temperature, water vapor, or ozone profiles if synergy is used even under clear sky conditions. The efficient use of simultaneous observations in various wavelength ranges makes it necessary to develop new retrieval strategies.

[4] In order to develop new approaches to perform satellite data fusion, it is necessary first to understand the basic concepts behind synergy. To illustrate the various types of synergy, it is a good strategy to use a simple schematic model. Then, a methodology needs to be put in place to measure the synergy. Since assimilation is the most widely used technique to merge data, the traditional information content analysis is the first candidate. To test the potential of this method, we apply it to selected atmospheric parameters and wavelength bands under specific instrument geometry for the MetOp-A satellite. This platform provides coincident observations in the infrared, with the Improved Atmospheric Sounding in the Infrared (IASI) instrument, and in the microwaves, with AMSU-A and Microwave Humidity Sounder (MHS). We concentrate on the major atmospheric parameters, namely temperature and water vapor profiles, for which the selected MetOp-A instruments are particularly sensitive.

[5] Section 2 presents various synergy mechanisms and illustrates them by using classical information content theory and a simple linear model. Section 3 introduces the simulated databases that are used in this study and by *Aires et al.* [2011]. The use of the classical information content analysis is illustrated using the MetOp application in section 4. Finally, conclusions and perspectives are drawn in section 5 and the companion paper is introduced.

## 2. Theoretical Considerations

### 2.1. Model

[6] Synergy (from the Greek syn-ergo, working together) refers to the phenomenon in which two or more discrete influences or agents acting together create an effect greater than that predicted by knowing only the separate effects of the individual agents.

[7] Let $f_1$ and $f_2$ be two geophysical variables, and $TB_1$ and $TB_2$ two measurements from two instruments. We suppose that the radiative transfer model can be linearized:

$$\begin{pmatrix} TB_1 \\ TB_2 \end{pmatrix} = A \cdot \begin{pmatrix} f_1 \\ f_2 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix} \quad (1)$$

where $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ is the linearized radiative transfer operator and $(\varepsilon_1,\varepsilon_1)$ is the measurement noise on $(TB_1,TB_2)$ that follows a Gaussian distribution with covariance matrix $S_\varepsilon = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$. We suppose that a priori information on $(f_1,f_2)$ exists and that its uncertainty is given by an unbiased Gaussian distribution with covariance matrix $S_f$. Of course, the following discussion can be extended to more geophysical variables, more observations or nonlinear models and non-Gaussian distributions.

[8] The solution of this simple linear and Gaussian inverse problem is given by:

$$f = f_g + \left(A^t \cdot S_\varepsilon{}^{-1} \cdot A + S_f{}^{-1}\right)^{-1} \cdot A^t \cdot S_\varepsilon{}^{-1} \cdot \left(F_\varepsilon - F_g\right) \quad (2)$$

where $f_g$ is the First Guess (FG), e.g., from a NWP model, $S_f$ is the FG covariance error, $A$ is the Jacobian matrix (i.e., first derivative of satellite observations with respect to geophysical variables $f$), $S_\varepsilon$ is the covariance of measurement errors, $F_\varepsilon$ are the satellite observations, and $F_g = A \cdot f_g$ [*Rodgers*, 1990]. The uncertainty on this solution is given by:

$$Q = \left(A^t \cdot S_\varepsilon^{-1} \cdot A + S_f^{-1}\right)^{-1} \quad (3)$$

This two-dimensional formulation of uncertainty is valid for the simultaneous use of both $TB_1$ and $TB_2$. For the use of only one satellite measurement, the formula has the same shape but is one-dimensional (i.e., with scalars instead of matrices).

[9] This simple theoretical radiative transfer and its associated retrieval model will allow illustration of the different types of synergies that can act in an observing system. In particular, in the context of remote sensing, we can identify four main types of synergy (Figure 1): additive (Figure 1a), unmixing (Figure 1b), indirect (Figure 1c), and denoising (Figure 1d). These four synergy cases are analyzed using this simple model in sections 2.2–2.6. This list is not exhaustive, and all the details cannot be treated in this paper (Figure 1b can also consider the case where a single measurement $TB_1$ is sensitive to two geophysical variables $f_1$ and $f_2$ but we will not discuss this case). Nonlinear synergy will also be discussed, but qualitatively only. We will consider the retrieval of $f_1$ when using only $TB_1$, only $TB_2$ or both of them simultaneously.

### 2.2. No Synergy

[10] For this example, let: $A = \begin{pmatrix} 0.8 & 0.0 \\ 0.0 & 0.9 \end{pmatrix}$ and suppose no correlation exists between the two geophysical variables, $cor(f_1, f_2) = 0$, with: $S_f = \begin{pmatrix} 3 & 0 \\ 0 & 4 \end{pmatrix}$.The coefficients 0.8 and

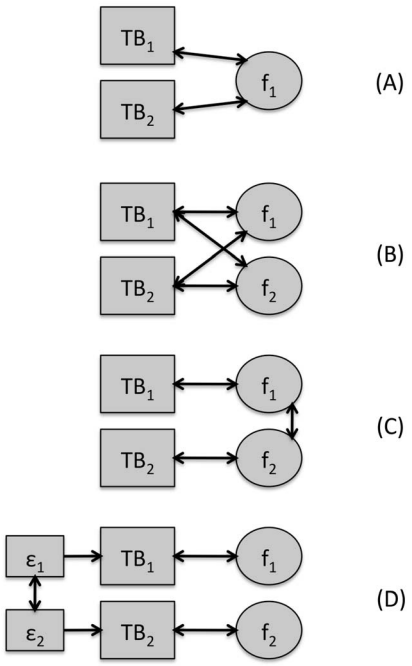


**Figure 1.** Four synthetic types of synergy: (a) additive synergy, (b) unmixing synergy, (c) indirect synergy, and (d) denoising synergy. $TB_1$ and $TB_2$ are the satellite observations, $\varepsilon_1$ and $\varepsilon_1$ are the corresponding instrument noise, and $V_1$ and $V_2$ are the geophysical variables to retrieve. Arrows represent dependencies of one variable to another, in the forward or inverse model. This scheme is valid for linear or nonlinear models.

0.9 in matrix $A$ are close to a sensitivity of one brightness temperature observations to an atmospheric temperature, $\frac{\partial TB}{\partial T}$. Using equation (3), it can be estimated that:

$$q_{f_1}^{TB_1} = \left(0.8 \cdot (1)^{-1} \cdot 0.8 + 3^{-1}\right)^{-1} = 1.0274$$

$$q_{f_2}^{TB_2} = \left(0.9 \cdot (2)^{-1} \cdot 0.9 + 4^{-1}\right)^{-1} = 1.5267$$

$$q_{(f_1,f_2)}^{(TB_1,TB_2)} = \begin{pmatrix} 1.0274 & 0 \\ 0 & 1.5267 \end{pmatrix}$$

where $q_f^{TB1}$ is the uncertainty on the retrieval of $f$ using $TB_1$ only and $q_{(f_1,f_2)}^{(TB_1,TB_2)}$ is the uncertainty matrix for the retrieval of ($f_1$,$f_2$) using ($TB_1$,$TB_2$). As can be seen in these computations, there is no improvement in the retrieval when using both measurements compared to using them independently (i.e., the retrieval uncertainty remains identical). Since there is no relationship between $TB_1$ and $f_2$ and between $TB_2$ and $f_1$, there is no additive or unmixing synergies. Since no correlation exists between $f_1$ and $f_2$, the indirect synergy does not apply either.

### 2.3. Additive Synergy

[11] Suppose now that both instruments provide direct information on $f_1$, so $A = \begin{pmatrix} 0.8 & 0 \\ 0.9 & 0 \end{pmatrix}$ but each one adds a contribution to the retrieval. Matrix $S_f$ is here just equal to 3. The uncertainty on the retrieval of $f_1$ using only $TB_1$, only $TB_2$ or both is given by:

$$q_{f_1}^{TB_1} = \left(0.8 \cdot (1)^{-1} \cdot 0.8 + 3^{-1}\right)^{-1} = 1.0274$$

$$q_{f_1}^{TB_2} = \left(0.9 \cdot (2)^{-1} \cdot 0.9 + 3^{-1}\right)^{-1} = 1.3544$$

$$q_{f_1}^{(TB_1,TB_2)} = \left(A^t \cdot (S_\varepsilon)^{-1} \cdot A + 3^{-1}\right)^{-1} = 0.7255$$

Even if the second measurement $TB_2$ has more noise than $TB_1$ and therefore is, when used independently, less precise, adding $TB_2$ helps reduce the uncertainty. The additive synergy is the direct consequence of the central limit theorem: using multiple measurements of the same variable reduces the uncertainty of the estimation.

[12] In Figure 2, the impact of the strength, $a_{21}$, i.e., the coefficient in $A$ linking $f_1$ and $TB_2$ on the retrieval accuracy, is measured. The retrieval accuracy using only $TB_1$ is constant and equal to 1.0274. For $a_{21} = 0$, the accuracy of 3 for the retrieval using only $TB_2$ is just the result of the a priori on $f_1$. As expected, the accuracy increases for increasing $a_{12}$. The synergy is always positive (i.e., factor bigger than 1), and $q_{f_1}^{(TB_1,TB_2)}$ is always lower than the best estimate from $TB_1$ or $TB_2$.

[13] The impact of instrument noise $\varepsilon_2$ in $TB_2$ is measured in Figure 3. When this noise is equal to zero, $TB_2$ gives a perfect retrieval, as expected. The retrieval from $TB_2$ only gets less accurate when $\varepsilon_2$ increases. Again, there is no impact in the accuracy level for the retrieval from $TB_1$ only. The synergy is always positive, and $q_{f_1}^{(TB_1,TB_2)}$ is always lower than the best estimate from $TB_1$ or $TB_2$. This proves that the additive synergy takes into account the instrument noise level in each instrument.

[14] As an example of such additive synergy, we can mention that microwave and infrared measurements can

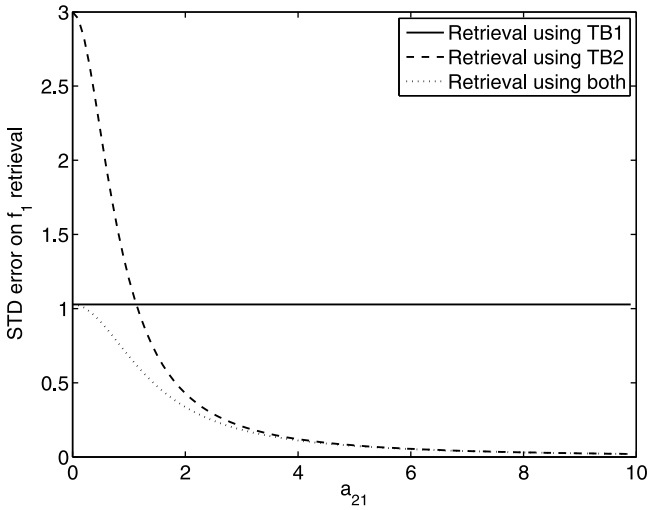


**Figure 2.** Sensitivity experiment for the additive synergy: standard deviation for the error in $f_1$ retrieval with respect to coefficient $a_{21}$ in model matrix $A$ (equation (1)) using only $TB_1$ (solid line), only $TB_2$ (dashed line), and both $TB_1$ and $TB_2$ (dotted line).

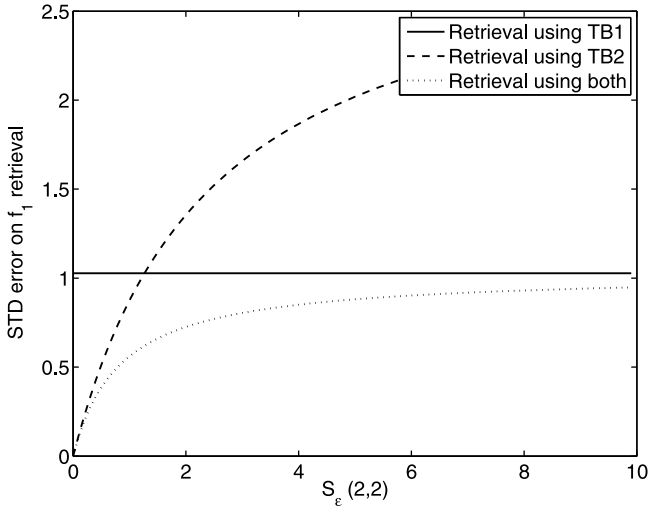


**Figure 3.** Sensitivity experiment for the additive synergy: standard deviation for the error in $f_1$ retrieval with respect to $TB_2$ noise level $\varepsilon_2$ using only $TB_1$ (solid line), only $TB_2$ (dashed line), and both $TB_1$ and $TB_2$ (dotted line).

provide information on temperature. The signal-to-noise ratio is different, the weighting functions peak at different altitudes, the vertical resolution can also be different but combination of information on the atmospheric temperature profile provides better estimation. Adding multisource information can transform an underconstrained problem into a well-posed problem.

### 2.4. Unmixing Synergy

[15] In this example, suppose $A = \begin{pmatrix} 0.8 & 0.2 \\ 0.4 & 0.9 \end{pmatrix}$ so both $TB_1$ and $TB_2$ have a sensitivity to $f_1$ but they also have a sensitivity on $f_2$ so there is a mixing of the $f$ signals on the satellite observations. Suppose also that $f_1$ and $f_2$ are uncorrelated and for example $S_f = \begin{pmatrix} 3 & 0 \\ 0 & 4 \end{pmatrix}$. The coefficients 3 and 4 in matrix $S_f$ are close to the uncertainties of an a priori atmospheric temperature. Again, using similar Bayesian estimation of the uncertainty using only $TB_1$, only $TB_2$ or both, we obtain:

$$q_{f_1}^{TB_1} = 1.1299$$

$$q_{f_1}^{TB_2} = 2.7483$$

$$q_{f_1}^{(TB_1,TB_2)} = 1.1274$$

Even if $TB_2$ has almost no information on $f_1$ (only a slight decrease in the uncertainty compared to the a priori equal to 3), we can observe a slight decrease (0.212%, from 1.1299 to 1.1274) of the retrieval uncertainty when using both $TB_1$ and $TB_2$. The impact of the synergy is not important in this case because the mixing between the variables is limited: the term $a_{12}$ in matrix $A$ that controls the impact of $f_2$ in measurement $TB_1$, at the origin of one mixing in the problem, is small, equal to 0.2. However, it is important to note that this decrease in retrieval uncertainty is not an additive synergy from the three sources of information ($TB_1$, $TB_2$, and the a priori on $f_1$) that, if simply estimated as a sum of information, would underestimate the real uncertainty. Furthermore, the retrieval error covariance matrix for $(f_1, f_2)$ will have off-diagonal elements (even if their a priori covariance matrix does not). So, if both are retrieved together, one gets a bit more information, namely, how the retrieved $f_1$ and $f_2$ are correlated.

[16] In Figure 4, the three terms $q_{f_1}^{TB_1}$, $q_{f_1}^{TB_2}$, $q_{f_1}^{(TB_1,TB_2)}$ are represented when varying parameter $a_{12}$. The retrieval uncertainty from the $TB_2$ is constant at 2.7483. The uncertainty from $TB_1$ increases when $a_{12}$ increases because its $f_1$ information gets corrupted by the influence of $f_2$. The unmixing gets more and more important when the mixing increases. For $a_{12} = 0.76$ (i.e., corresponding to a rather large mixing), $q_{f_1}^{TB_1} = 1.8987$ and $q_{f_1}^{(TB_1,TB_2)} = 1.8196$ which represents a synergy factor of more than 4%. As a consequence, if the retrieval methodology is able to untangle the information on $TB$s that mixes the contribution of the geophysical parameters $f_1$ and $f_2$, the estimation can benefit from this unmixing synergy. The higher the mixing is, the more beneficial the synergy becomes. This unmixing process is the equivalent of introducing new constrains in the inverse problem.

[17] An example of such synergy is given by microwave and infrared measurements that provide information on surface temperature, but this signal is mixed with information on surface emissivity. By using a combination of microwave and infrared observations, it is easier to assess the respective contribution of the surface temperature and emissivities and therefore to untangle the information mixture. The partial "decontamination" of the emissivity signal allows for a more accurate retrieval of surface temperature [*Aires et al.*, 2001].

### 2.5. Indirect Synergy

[18] In this case, the direct model is driven by the matrix $A = \begin{pmatrix} 0.8 & 0.0 \\ 0.0 & 0.9 \end{pmatrix}$ so $TB_1$ gives information only on $f_1$ and

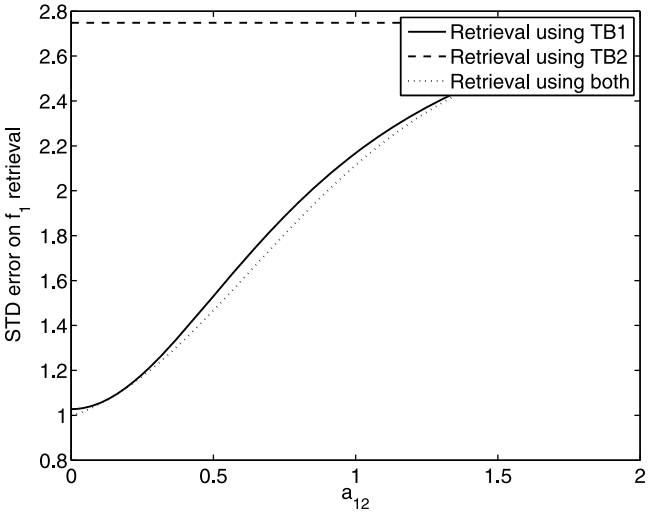


**Figure 4.** Sensitivity experiment for the unmixing synergy: standard deviation for the error in $f_1$ retrieval with respect to coefficient $a_{12}$ in model matrix $A$ (equation (1)) using only $TB_1$ (solid line), only $TB_2$ (dashed line), and both $TB_1$ and $TB_2$ (dotted line). Parameter $a_{12}$ controls the impact of state variable $f_2$ on $TB_1$; the higher $a_{12}$, the higher the mixing.

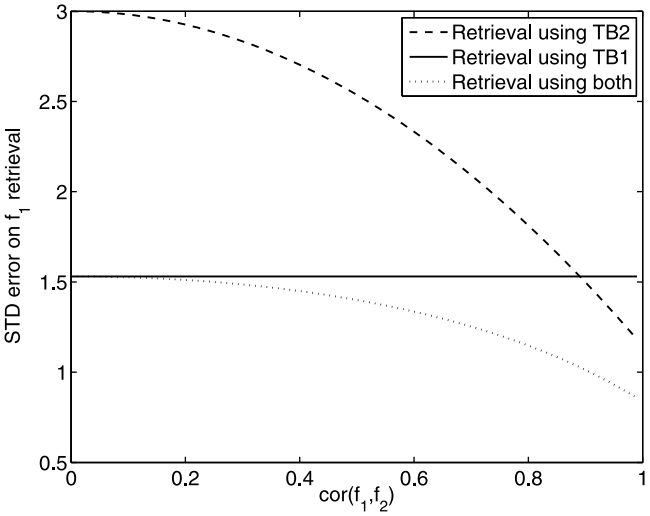


**Figure 5.** Sensitivity experiment for the indirect synergy: standard deviation for the error in $f_1$ retrieval with respect to correlation coefficient $C(f_1,f_2)$ using only $TB_1$ (solid line), only $TB_2$ (dashed line), and both $TB_1$ and $TB_2$ (dotted line).

$TB_2$ only on $f_2$ but it is supposed that $f_1$ and $f_2$ are correlated to the 0.5 level: $S_f = \begin{pmatrix} 3.00 & 1.73 \\ 1.73 & 4.00 \end{pmatrix}$. The retrieval of $f_1$ using only $TB_2$ makes no sense. The uncertainty on the retrieval of $f_1$ using only $TB_1$ or both measurements ($TB_1$,$TB_2$) is given by:

$$q_{f_1}^{TB_1} = 1.0274$$

$$q_{f_1}^{(TB_1,TB_2)} = 0.9664$$

The indirect information $TB_2 \rightarrow f_2 \rightarrow f_1$ allows for the improvement by still 6% of the estimation of $f_1$. We refer, in this case, to an "indirect synergy."

[19] Channels from different infrared or microwave instruments provide information on temperature. Since temperature and water vapor atmospheric profiles are related physically and statistically, it is possible to use this dependency inside the retrieval process. This is done routinely in assimilation by using the temperature/water vapor covariance matrix. This can also be used in a neural network retrieval [*Aires et al.*, 2002; *Karbou et al.*, 2005].

[20] In Figure 5, the two terms $q_{f_1}^{TB_1}$ and $q_{f_1}^{(TB_1,TB_2)}$ representing the retrieval accuracy for $TB_1$ only or for both $TB$s is represented for increasing correlation $C(f_1,f_2)$ that impacts matrix $S_f$. When the correlation $C(f_1,f_2) = 0$ the information comes only from $TB_1$ and none from $TB_2$ so $q_{f_1}^{TB_1} = q_{f_1}^{(TB_1,TB_2)} = 1.0274$. The indirect synergy appears when this correlation increases and $q_{f_1}^{(TB_1,TB_2)}$ gets much lower than $q_{f_1}^{TB_1}$ and its minimum equals to 0.6 for $C(f_1,f_2) = 1$ representing more than 40% synergy.

### 2.6. Denoising Synergy

[21] The direct model is driven by the matrix $A = \begin{pmatrix} 0.8 & 0.0 \\ 0.0 & 0.9 \end{pmatrix}$ so $TB_1$ gives information only on $f_1$ and $TB_2$ only on $f_2$. Suppose $f_1$ and $f_2$ are uncorrelated: $S_f = \begin{pmatrix} 3 & 0 \\ 0 & 4 \end{pmatrix}$ but we consider a correlation $cor(\varepsilon_1,\varepsilon_2) = 0.5$ between the two instrument noises.

[22] The retrieval of $f_1$ using only $TB_2$ or the retrieval of $f_2$ using only $TB_1$ still makes no sense. The uncertainty of the retrieval of $f_1$ using only $TB_1$, the retrieval of $f_2$ using only $TB_2$ and the uncertainty on the retrieval of ($f_1$, $f_2$) by using both measurements ($TB_1$,$TB_2$) is given by:

$$q_{f_1}^{TB_1} = 1.0274$$

$$q_{f_2}^{TB_2} = 1.5267$$

$$q_{(f_1,f_1)}^{(TB_1,TB_2)} = \begin{pmatrix} 0.9608 & 0.41128 \\ 0.41128 & 1.4432 \end{pmatrix}$$

There is a strong synergy: when using both measurements, the synergy (i.e., decrease of the uncertainty when using both measurements together) is 6.5% for the estimation of $f_1$, and 5.47% for the estimation of $f_2$. Note also, when looking at matrix $q_{(f_1,f_1)}^{(TB_1,TB_2)}$, that there is an advantage to retrieve simultaneously $f_1$ and $f_2$: the correlation on the retrieval uncertainty can be directly estimated. The correlation is here $0.41128/\sqrt{0.9608*1.4432} = 0.3493 \simeq 35\%$. This type of uncertainty correlation structure can be provided by most of the inversion algorithms. This is an extremely valuable information that can be exploited when the retrieval is used, for example, in a variational assimilation scheme.

[23] In Figure 6, the sensitivity of the synergy is measured with respect to the correlation between instrument noise measurements. The constant level for the retrieval uncertainty is represented when $TB_1$ and $TB_2$ are used independently. When the correlation $cor(\varepsilon_1,\varepsilon_2)$ increases, this uncertainty decreases significantly, for the retrieval of $f_1$ and $f_2$. This type of synergy is referred to as a "denoising synergy." An

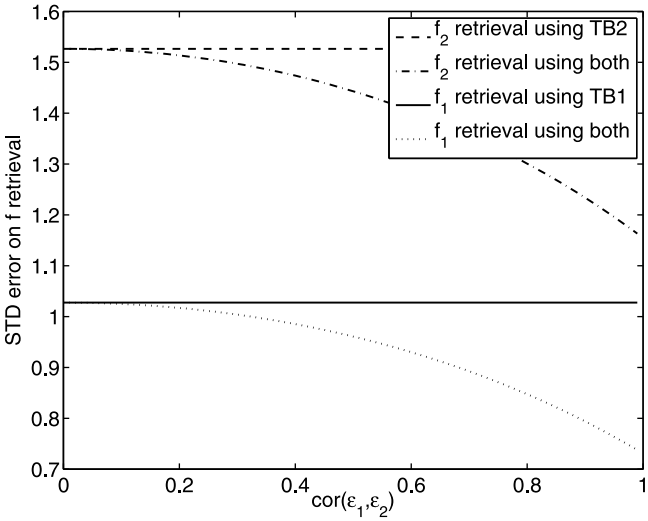


**Figure 6.** Sensitivity experiment for the denoising synergy: standard deviation for the error in $f$ retrieval with respect to correlation coefficient $C(\varepsilon_1,\varepsilon_2)$ for $f_2$ retrieval using only $TB_2$ (dashed line), $f_2$ retrieval using only both $TB_1$ and $TB_2$ (dot-dashed line), $f_1$ retrieval using only $TB_1$ (continuous line), and $f_1$ retrieval using both $TB_1$ and $TB_2$ (dotted line).

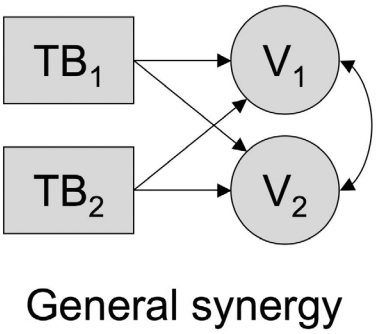


**Figure 7.** General synthetic synergy scheme: $TB_1$ and $TB_2$ are the satellite observations, and $V_1$ and $V_2$ are the geophysical variables to retrieve. This scheme is valid for linear or nonlinear models.

example of a denoising situation can be mentioned here because it will be pertinent in the following: $TB_1$ and $TB_2$ can be two consecutive channels in a IASI spectrum. Since the IASI spectrum is a convolution by an instrument function, the instrumental noise is correlated from one channel to another and of course, there input error correlation have an impact on the retrieval.

### 2.7. Nonlinear Synergy

[24] We will not enter into details in this section: it would require the introduction of a nonlinear extension of the model in equation (1) and the development of a new experimental procedure to measure synergy since equation (3) would not be valid anymore. We introduce here two points that illustrate nonlinear difficulties that can be solved when synergy is used.

[25] The first of these points concerns saturation effects. Consider a model where $TB_1$ and $TB_2$ are two satellite observations related to a geophysical variable $f$ and suppose that the instrument providing these measurements saturate (i.e., no significant response to an increase in $f$ after some threshold). $TB_1$ and $TB_2$ can saturate at a different threshold and as a consequence, each measurement would give an information on $f$ but for a different range or variability. When used together, $TB_1$ and $TB_2$ can characterize $f$ in a range which is the sum of the two initial ranges.

[26] Another nonlinear model difficulty occurs when interactions act in the measurement (this is a particular case of unmixing synergy). Let us consider the following model:

$$\begin{pmatrix} TB_1 \\ TB_2 \end{pmatrix} = \begin{pmatrix} a_1 & 0 \\ 0 & a_2 \end{pmatrix} \cdot \begin{pmatrix} f_1 \cdot f_2 \\ f_2 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix}$$

It is clear that $TB_1$ only cannot allow the retrieval of $f_1$. This problem can only be solved by using both $TB_1$ and $TB_2$. An example of such problem concerns the simultaneous retrieval of surface temperature and emissivity, already mentioned: in the microwave domain, for surface sensitive channels, the satellite observation can be approximated by the product of these two geophysical variables and disentangling them requires the combination of multiple satellite observations at different frequencies.

[27] It is important to point out that in real world applications, it is difficult to identify cases where only one single type of synergy exists. For example, where there is correlation between the variables to retrieve $f_1$ and $f_2$, it is most likely that there is mixing in the $A$ direct model: the correlations in the geophysical space can introduce "indirect correlations" between observations and geophysical variables. As a consequence, an observation can be correlated to multiple geophysical variables even if it is physically sensitive to only one of them, in terms of radiative transfer function. Since most of the inversion algorithms use this type of correlation information, it is convenient to consider these "indirect correlations." Figure 7 represents a generalization of some of the basic synergy processes. This type of configuration is closer to what can be found in nature in general, and in remote sensing problems in particular.

## 3. Geophysical and Satellite Observation Databases

[28] Most retrieval techniques require a data set of surface and atmosphere variables (e.g., surface and atmospheric temperature, water vapor and ozone profiles) representing the natural meteorological/climatic variability, with corresponding satellite observations (real or simulated). Here we use the analysis data set from ECMWF for most of the geophysical variables needed for radiative transfer simulations. Associated simulations will be performed using the radiative transfer models for MetOp-A instruments: AMSU-A, MHS and IASI. The corresponding synthetic data set will represent ten thousand situations (i.e., a sample of the geophysical database corresponding to particular surface and atmospheric conditions), enough to calibrate the inversion techniques, and in particular to train the Neural Network (NN) models. The global data set will be divided into a training database (to calibrate the inversion) and a testing database (to evaluate the generalization capacities of the inversion models).

[29] In this application, only cloud-free situations over the ocean are considered, with nadir viewing geometry.

### 3.1. ECMWF Operational Analyses

[30] The atmospheric profiles and surface properties from the 6-hourly operational global analyses from the Integrated Forecasting System (IFS) of the European Center for Medium Range Forecasting (ECMWF) [*Simmons and Gibson*, 2000] are at the origin of the data sets that will be constructed in this section. In order to run accurate radiative transfer simulations, the following information is kept: the temperature, water vapor (relative humidity in % hereafter) and ozone profiles on 43 pressure levels ranging from 1000 to 1 hPa (these levels have been interpolated for the initial 21 levels in order to be used with the RTTOV code) and surface properties such as the 10 m horizontal wind, 2 m pressure and temperature and surface temperature.

[31] Selecting cloud-free oceanic cases over a year yields on the order of a few million atmospheric and surface situations.

### 3.2. High-Dimensional Sampling Procedure

[32] Retrieval algorithms cannot handle this huge amount of data. In order to reduce the size of the previous geophysical data set while keeping its spatial and temporal variability, a sampling procedure is required. This sampling should work efficiently in high-dimensional space, i.e., the

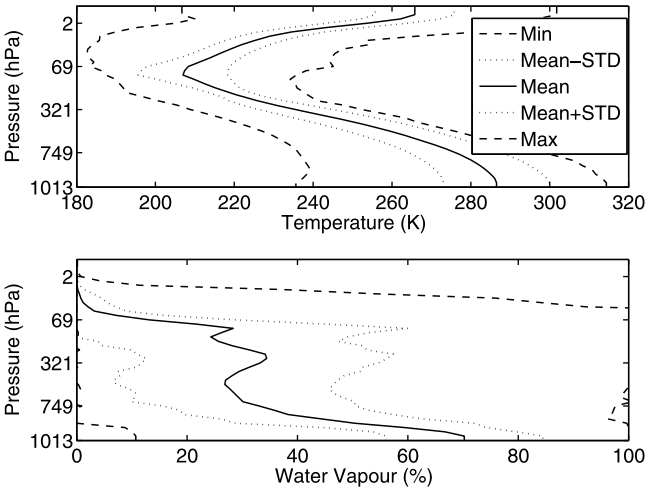


**Figure 8.** Statistics for the atmospheric data set. Minimum, maximum, and mean atmospheric profiles are represented for (top) temperature and (bottom) water vapor.

space of the geophysical variables which represents $3 \times 43 = 129$ (three profiles: temperature, $T$, water vapor, $H_2O$, and ozone, $O^3$, over 43 atmospheric layers), this defines the *GEO* space. Furthermore, since we are interested in retrieval algorithms exploiting synergy, the sampling procedure should be able to process truly multivariate samples. In the work by *Aires and Prigent* [2007], a clustering approach was proposed to perform this sampling. A clustering algorithm is a statistical method that extracts, from a large database, a set of prototypes that describe, as well as possible, the variability in the original data set.

[33] A major advantage of this clustering technique is that the resulting distributions are close to the original one. This means that the highly populated parts of the state space would be represented by a large number of prototypes, and that the low-population density parts of the state space will have less prototypes. This is not the case, for example, for the uniform sampling that obtains by definition more "uniform" distributions. Obtaining realistic distribution can be an advantage (a uniform sampling can also be preferred, depending on the application, the extreme events are favored and this can be more efficient for the learning of retrieval scheme): for example to perform Bayesian statistics, the distributions must respect the natural variability. Another advantage of clustering is that the number of clusters $K$ is defined a priori, contrarily to the uniform sampling approach.

[34] The $K$ means algorithm [*Lloyd*, 1992] is an example of clustering method and it is selected here to sample a large and high-dimensional data set (other clustering algorithm could be used). This clustering method has been used in the atmospheric science disciplines: for example, in the work by *Jakob et al.* [2005], $K$ means were used to relate radiative, cloud, and thermodynamic properties. The $K$ means algorithm has also been used to validate GCM cloud properties compared to satellite observations [*Chéruy and Aires*, 2009]. In the work by *Aires and Prigent* [2007], this tool was used to construct atmospheric databases. The $K$ means algorithm steps are simple:

[35] 1. First, $K$ prototypes are selected in the data space to be clustered (*GEO* space in our case). They are generally randomly chosen.

[36] 2. Assign each sample of the data set to its closest prototype by using the data distance. This cluster allocation determines $K$ clusters of points.

[37] 3. When all samples have been assigned, calculate the mean for each of the $K$ clusters. These cluster centers become the new prototypes. To add stability, a "learning rate" can be adopted so that the new mean is a linear combination of the previous mean and its new estimate [*Moody and Darken*, 1989].

[38] 4. Repeat steps 2 and 3 until the convergence is reached. A criterion checks this convergence: In our case, the training phase is stopped when the relative change in the prototypes is small. This is done by monitoring the relative change curves.

[39] The $K$ means clustering is sensitive to the metric being used. Few possible distances can be used (Euclidean, min or max, Mahalanobis), each one with its own advantages and inconveniences. The Euclidean distance is preferred here because the retrieval methods that will be described in section 4 use the least square criterion.

[40] Clustering approaches are time consuming when used to sample databases from large and high-dimensional data sets [*Aires and Prigent*, 2007]. In order to use this technique for this particular application, we designed a novel hierarchical clustering strategy: In a first stage, the clustering is used to extract 100 first-generation prototypes. Each prototype is associated to a cluster of data in the original data set. In the second stage, a new clustering is performed in each of the 100 clusters of data in order to find 100 second-generation prototypes. With this approach, $100 \times 100 = 10.000$ prototypes are extracted from the original data set and the procedure is very computationally efficient. (Furthermore, the fact that the extracted database is structured hierarchically can be a very interesting feature. As an example, when a Bayesian algorithm uses a precipitation database to retrieve rain, for each new observation a search for the closest situations in the database needs to be done, and this step can be extremely time consuming. With the hierarchical feature, the search for the closest situations could first identify the closest first-generation prototype, and then search in its associated second-generation prototypes for the closest situations.)

[41] Some statistics on the atmospheric temperature and water vapor profiles are shown in Figure 8. The range of variability in this geophysical space covers all types of situations and is adequate for the training of retrieval techniques since by definition, all the atmospheric situations from one year of ECMWF analysis can be represented by one of the clusters of the database. Another quality check of the generated database is to verify the location of the extracted samples. Note that the clustering does not use any geographical information: the sampling is performed using only temperature and water vapor information. In Figure 9, the atmospheric temperature of the near-surface atmospheric layer is represented. This shows the ocean areas covered by our data set. The longitude and latitudinal coverage is good. The coastal areas seem to be more represented, the reason is that the coastal situations are submitted to the continental atmospheric behavior that is more variable and complex

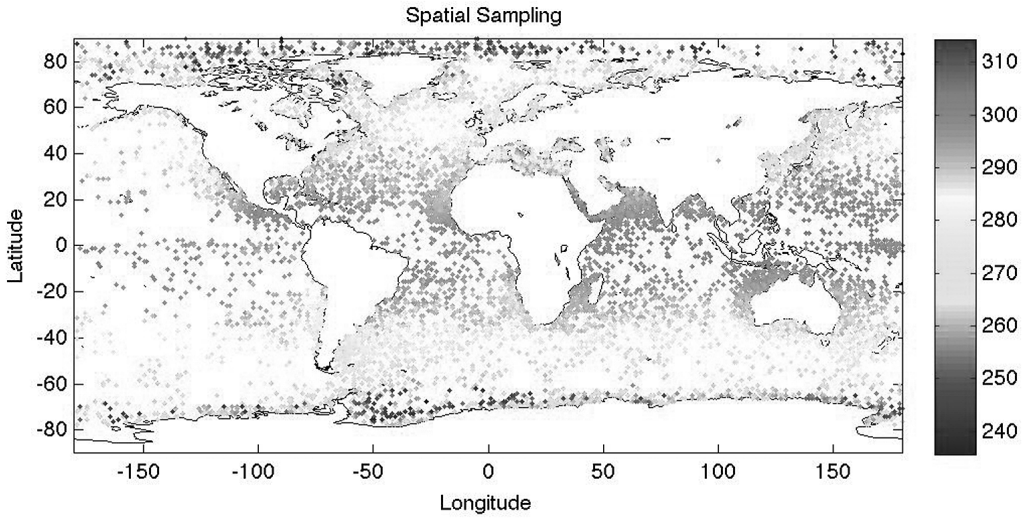


**Figure 9.** Spatial location of the 10,000 atmospheric situations over the ocean used to train and test the retrieval methods. The grey shades represent the temperature (K) at the surface layer for illustration purposes.

than the situations in open sea. Overall, this check of the specifically designed database confirms that it is as a good calibration database for retrieval algorithms.

### 3.3. Satellite Instruments

[42] Launched on 19 October 2006, MetOp is Europe's first polar-orbiting satellite dedicated to operational meteorology. It is a series of three satellites to be launched sequentially over 14 years, forming the space segment of EUMETSAT's Polar System (EPS). MetOp carries a set of "heritage" instruments provided by the United States and a new generation of European instruments that offer improved remote sensing capabilities to both meteorologists and climatologists. The new instruments increase the accuracy of temperature humidity measurements, wind speed and direction, and atmospheric ozone profiles. MetOp flies in a polar orbit corresponding to local "morning."

[43] In this study, simulated observations from the following instruments will be used. The Advanced Microwave Sounding Unit-A (AMSU-A) measures the oxygen band between 50 and 60 GHz, for the retrieval of atmospheric temperature profiles [*Mo*, 1996]. It is a cross-track scanning radiometer, with ± 48.3° from nadir with a total of 30 Earth fields of view of 3.3° per scan line, providing a nominal spatial resolution of 48 km at nadir. The instrument completes one scan every 8 s. The swath width is approximately 2000 km. AMSU-A is divided into two separate modules: (1) AMSU Module A1 with channels 3 to 15 (12 sounding channels in the 55 GHz $O_2$ band and one at the 89 GHz window) and (2) AMSU Module A2 with channels 1 and 2 at 23.8 and 31.4 GHz. AMSU-A is used in conjunction with the High-resolution Infrared Sounder instrument to estimate the global atmospheric temperature and humidity profiles from the surface to the upper stratosphere ($\simeq$50 km). AMSU-A measurements also provide precipitation and surface information including snow cover, sea ice concentration and soil moisture.

[44] The Microwave Humidity Sounder (MHS) is designed to measure the atmospheric water vapor profile, with 3 channels in the $H_2O$ line at 183.31 GHz plus two window channels at 89 and 150 GHz [*Hewison and Saunders*, 1996]. MHS scans the Earth from left to right, in a vertical plane. Each swath is made up of 90 contiguous individual pixels sampled every 2.67 s. The scan is also synchronized with the AMSU-A1 and AMSU-A2 instruments.

[45] The Infrared Atmospheric Sounding Interferometer (IASI) is a state-of-the-art Fourier transform spectrometer based on a Michelson interferometer coupled to an integrated imaging system that observes and measures infrared radiation emitted from the Earth [*Chalon et al.*, 2001]. It has been developed by the French space agency CNES. The optical interferometry process offers fine spectral samplings of the atmosphere in the infrared band between the 3.2 and 15.5 microns representing 8461 channels. This enables the instrument to retrieve temperature and water vapor profiles in the troposphere and the lower stratosphere, as well as measure concentrations of ozone, carbon monoxide, methane and other compounds. For optimum operation, the IASI measurement cycle is synchronized with that of the AMSU-A1 and AMSU-A2. This instrument was designed to reach accuracies of 1 K in temperature and 10% in water vapor with vertical resolutions of 1 km and 2 km, respectively, for cloud-free scenes.

[46] IASI, together with AMSU-A and MHS, has led to important improvements in the accuracy of remotely sensed temperature and humidity profiles and ozone amount.

### 3.4. Radiative Transfer Simulations

[47] The RTTOV-8.7 radiative transfer model is used to simulate the AMSU-A and MHS channels from the description of the atmosphere given by the geophysical database described in sections 3.1 and 3.2. This model, originally developed at ECMWF [*Eyre*, 1991] and now supported by the EUMETSAT NWP-SAF (Numerical Weather Prediction-Satellite Application Facility) [*Saunders et al.*, 1999; *Matricardi et al.*, 2004], provides rapid simulations of radiances for satellite infrared and microwave

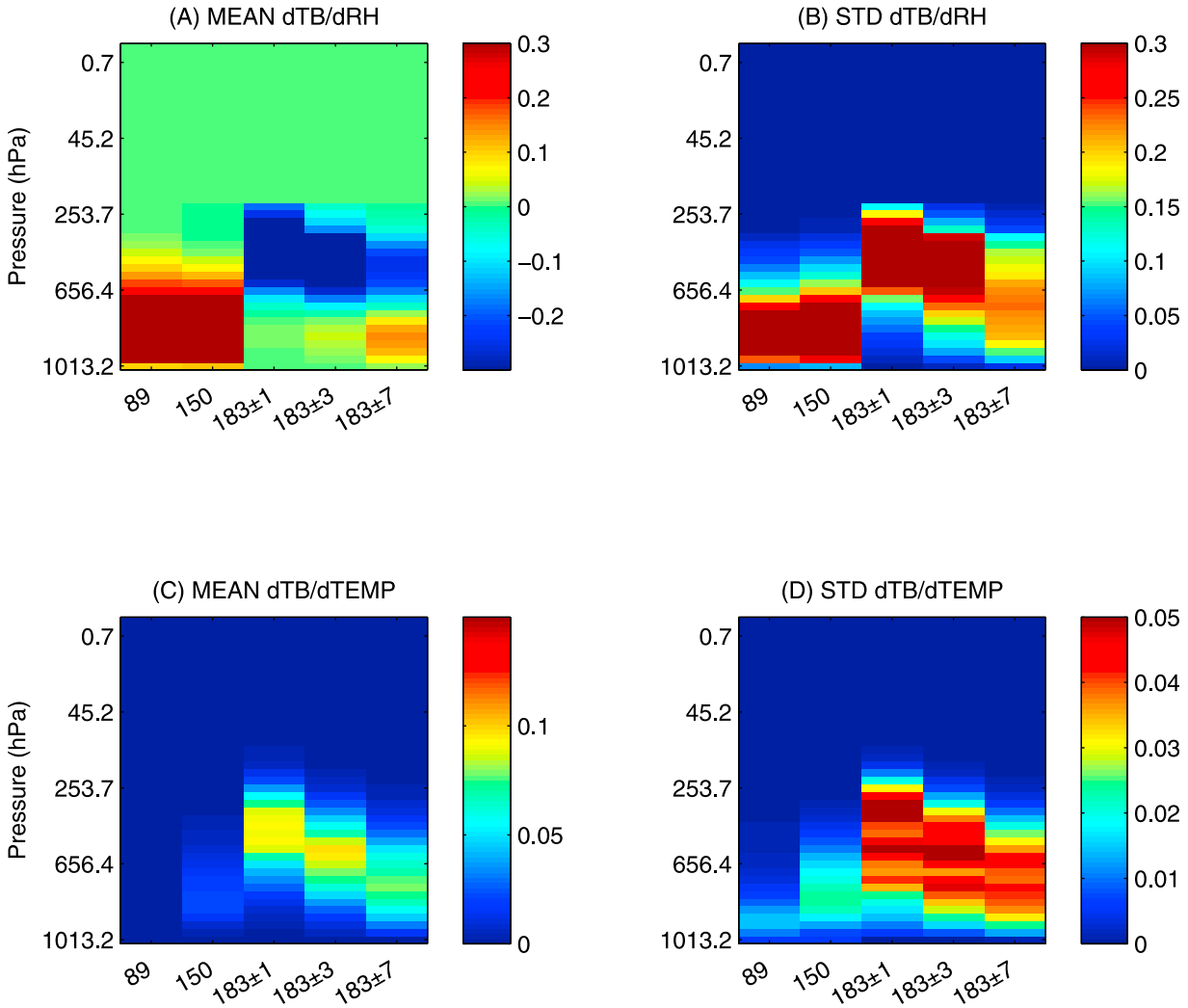


**Figure 10.** Mean and standard deviation of the Jacobian of the MHS observations with respect to (a and b) water vapor and (c and d) temperature over the ocean.

radiometers for a given atmospheric state vector. Over ocean, the emissivities are computed by the FASTEM-3 [*Deblonde and English*, 2001] surface emissivity model.

[48] The Automatized Atmospheric Absorption Atlas (4A) is a fast and accurate Radiative Transfer Model (RTM) for the infrared codeveloped by LMD and Noveltis with the support of CNES [*Scott and Chédin*, 1981]. The 4A computes transmittance and radiance, using a comprehensive database (atlases) of monochromatic optical thicknesses for up to 43 atmospheric molecular species. Precomputed once and for all, the atlases are derived from a line-by-line and layer-by-layer model, Stransac (i.e., a line-by-line radiative transfer) [*Scott*, 1974], with up-to-date physics. It uses spectroscopy from the GEISA spectral line data catalog but other spectroscopy data banks can be used as well. 4A/OP (OPerational) is chosen here to perform the IASI simulations because it is the reference RTM for the EUMETSAT IASI level 1 cal/val and level 1 operational processing.

[49] Both of these RTM provide analytical Jacobians, i.e., first derivative of the satellite observations with respect to the geophysical RTM inputs: $\frac{TB}{\partial x}$ for example. Brightness temperatures (K) are used in this study instead of radiances because, in this way, ranges of variability of each channels can be compared more easily.

## 4. Theoretical Information Content

[50] The most widely used technique exploiting synergy among Earth observation instruments is, without any doubt, the assimilation [*Kalnay*, 2003]: a wide spectrum of visible, infrared and microwave satellite observations are combined with model forecasts and in situ measurements to better characterize and predict the state of the atmosphere, continental surfaces or oceans. In association to the assimilation technique, various tools have been designed to estimate the theoretical quality of retrievals. Since we are interested in this study in the synergy, it is essential to test this approach first. Assimilation and information content analysis share the same theoretical hypothesis: Gaussian character of the stochastic variables, linearization around the First Guess (FG), same observation, radiative transfer and a priori uncertainties. See *Tarantola* [1987] for a general textbook

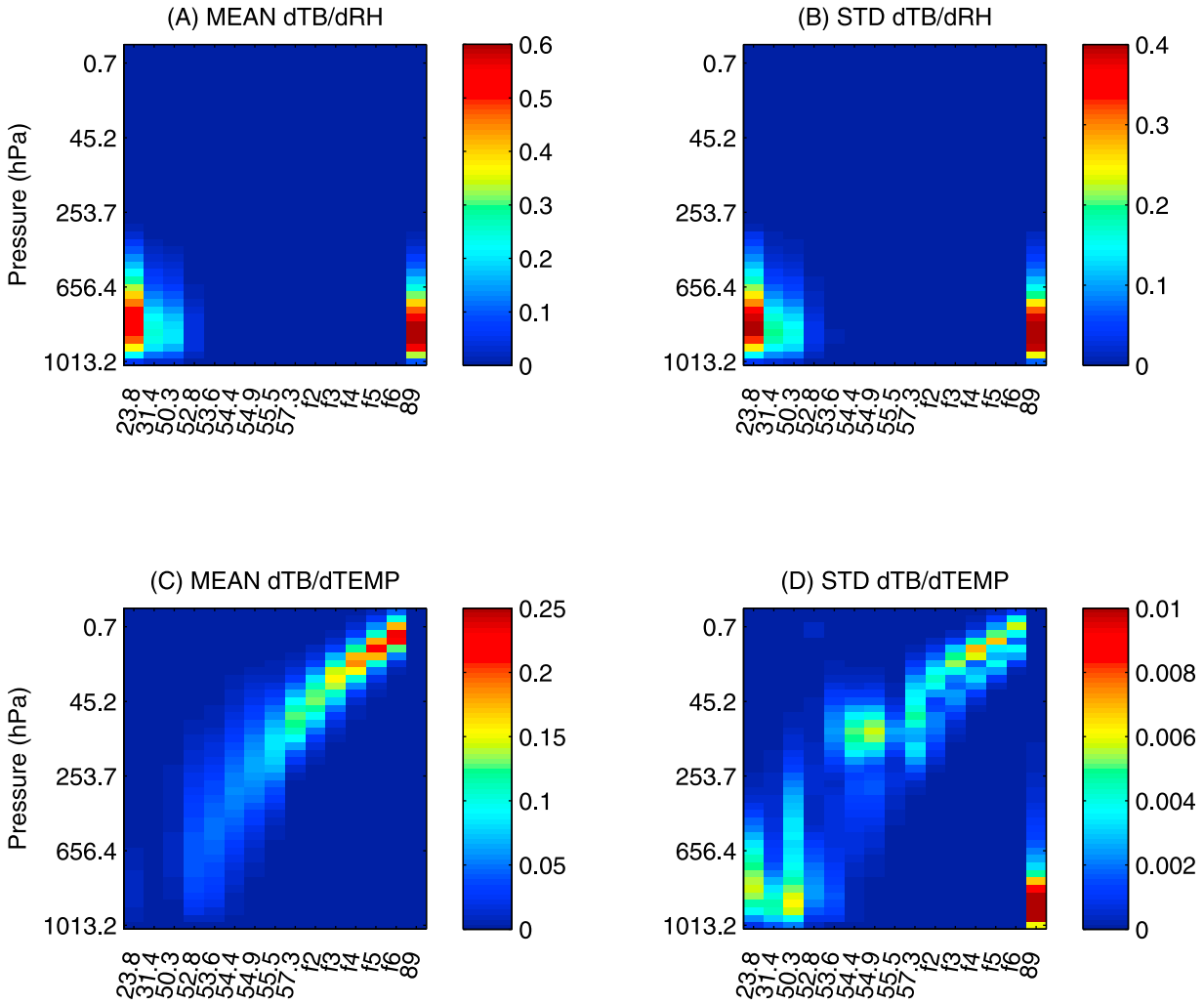


**Figure 11.** Mean and standard deviation of the Jacobian of the AMSU-A observations with respect to (a and b) water vapor and (c and d) temperature over the ocean.

on inverse problems and *Rodgers* [2000] for dedicated remote sensing theory.

[51] One of the most useful formula from the information content analysis is given in equation (3) that was previously used in section 2. This formula is a direct consequence of the Bayes rule and allows for the estimation of retrieval uncertainties. This quantity depends only on the Jacobians of the RTM, the observational noise and the a priori uncertainty. The central questions of this section are the following: (1) Are these estimations reliable? (2) Can this tool be used to measure the synergy between measurements?

### 4.1. RTM Jacobians

[52] In order to use equation (3), the Jacobian of the RTM needs to be estimated for the three instruments that are considered here, namely MHS, AMSU-A and IASI. Both RTTOV and 4A radiative transfer models (section 3.4) provide analytical Jacobians. 4A provides Jacobians in specific humidity. Relative humidity Jacobians are preferred here because the retrieval is performed on this unit. As a consequence, and for practical reasons, the Jacobians are estimated using RTM simulations on perturbed input profiles. The perturbations are chosen to be 1 K for temperature and 10 % for relative humidity.

[53] Figures 10, 11, and 12 represent the temperature and relative humidity mean Jacobians, together with the standard deviation of these Jacobians for MHS channels, ASMU-A, and IASI, respectively. The mean and standard deviation of Jacobians are estimated using a diverse set of 100 atmospheric situations from the database described in section 3.2. It appears that MHS instrument is as expected more sensitive to changes in relative humidity than to changes in temperature (±0.4 K compared to 0.1 K on average). AMSU-A provides higher magnitudes of Jacobians for both temperature and water vapor but MHS provides good sensitivity in the lower atmosphere (below 300mb) for temperature and a better vertical coverage for water vapor. Note nevertheless that, for the temperature, the standard deviation of the MHS Jacobians is very large compare to the one from AMSU-A, meaning that other factors, likely relative humidity in this case, also contaminate the signal. The magnitudes of Jacobians are comparable for the three instruments, including the infrared from IASI. However, the

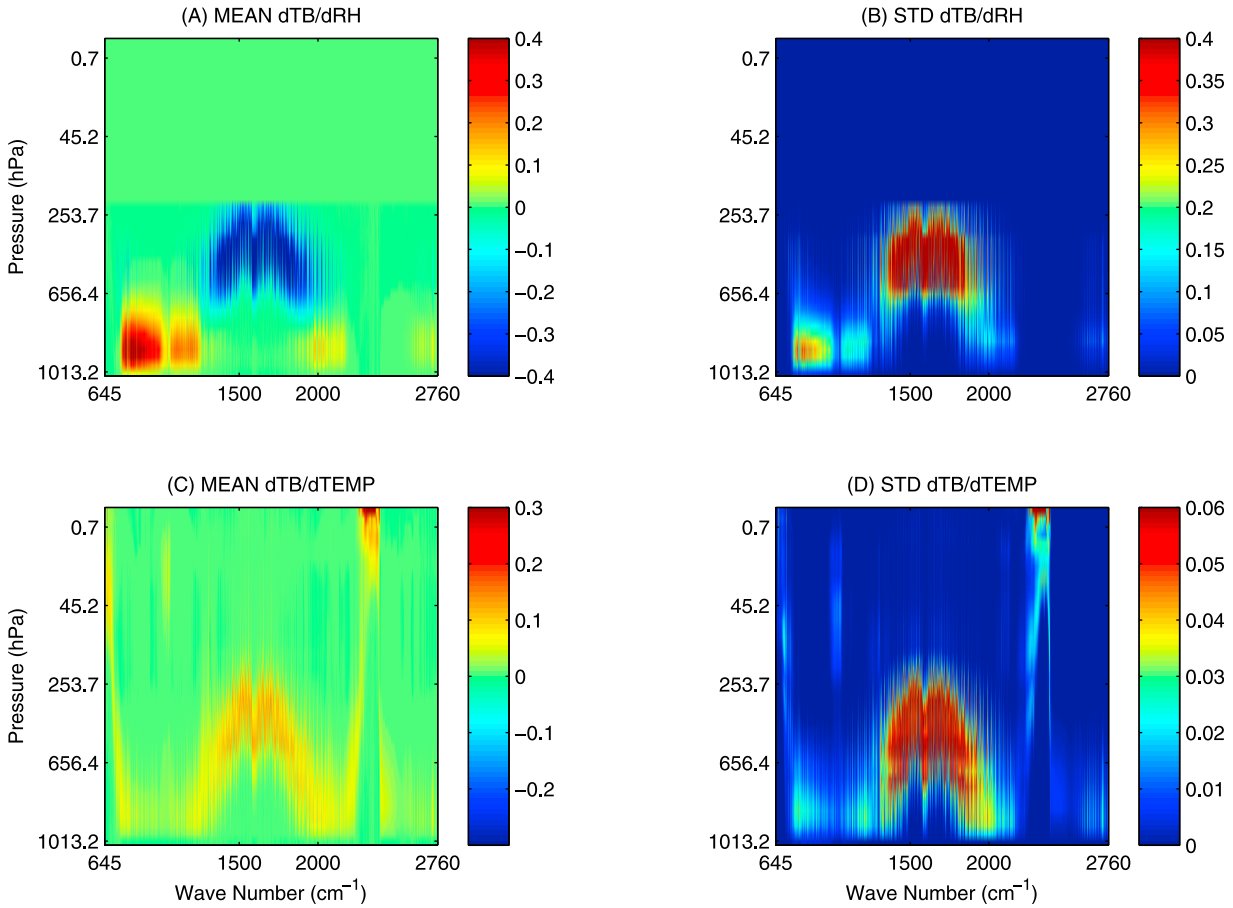


**Figure 12.** Mean and standard deviation of the Jacobian of the IASI observations with respect to (a and b) water vapor and (c and d) temperature over the ocean.

vertical resolution for IASI is higher for water vapor in the upper troposphere, and for the temperature for the whole troposphere. Furthermore, the IASI instrument possesses a lot of channels, this has a lot of consequences for the retrievals (i.e., computation time for the retrieval, necessity to perform a dimension reduction on the observed spectra, redundancy considerations for the denoising, etc.). It should be noted that the water vapor Jacobians are positive for the lower troposphere but are negative above 500 hPa. This is true for the three instruments. It should also be noted that, for temperature, the standard deviation (STD) of Jacobians is much lower than the mean Jacobian values but that for water vapor, the STD is very high compared to the mean Jacobian. The Jacobian sign can even change from a situation to another. It is expected that this will have a strong impact on the retrieval uncertainty estimates.

### 4.2. A Priori Information

[54] An essential element of the information content estimation of the retrieval uncertainties is the a priori information: first, the a priori is an additional "virtual" observation [*Rodgers*, 1990, 2000] which improves the retrieval. Second, equation (3) supposes that the RTM, that is strongly nonlinear by nature, is linearized around a FG. Third, the introduction of this additional information regularizes the inverse problem and without it, the inversion matrix of the first term in equation (3) is often ill conditioned.

[55] Covariance matrices $S_f$ with a simple structure can be used and are often chosen to be diagonal (or tridiagonal, i.e., a matrix that has nonzero elements only in the main diagonal, the first diagonal below this, and the first diagonal above the main diagonal), with some specified variance error for the a priori information. Another proxy for the a priori covariance matrix is produced in operational meteorological centers by using the departure of the analysis with respect to the prediction [*Rabier et al.*, 1998]. This can be considered as being the FG error if the analysis is the true solution.

[56] In order to obtain a realistic a priori covariance matrix, the data set of 10000 atmospheric situations presented in section 3.2 is divided in two data sets, $\mathcal{D}_1$ and $\mathcal{D}_2$. The first data set, $\mathcal{D}_1$, is composed of 2000 situations that are kept to perform a pattern recognition; the other data set, $\mathcal{D}_2$, is composed of 8000 situations and is used to test the pattern recognition. Each situation in $\mathcal{D}_2$ is described by its geophysical variables and its associated *TB*s (i.e., the satellite observations). The TBs are compared to each *TB*s in $\mathcal{D}_1$ using a regular Euclidian distance on MHS, AMSU-A and IASI *TB*s. The geophysical profile of the closest situation in $\mathcal{D}_1$ is chosen to be the FG for the retrieval. (This pattern recognition procedure is a very simple retrieval

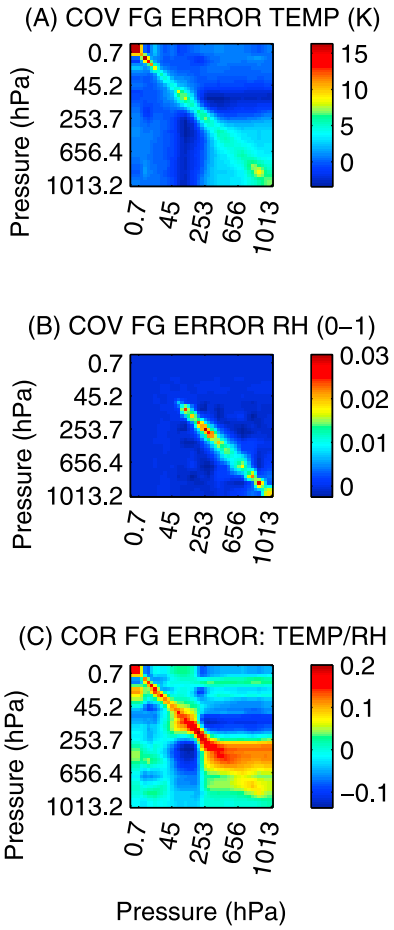


**Figure 13.** The covariance of the a priori errors for (a) temperature and (b) water vapor. (c) The correlation matrix between the a priori errors on temperature and water vapor.

scheme that will be tested by *Aires et al.* [2011].) The difference between the real situation in $\mathcal{D}_2$ and the FG in $\mathcal{D}_1$ is used to built the a priori statistics, in particular the $S_f$ covariance matrix. It should be noted that since this FG is dependent on the observations, it is not a regular a priori but this is nonetheless a good estimation of it: the improvement resulting from the use of the observations is compensated, in practice, by better a priori from a prediction model. In order to obtain realistic statistics, synthetic observational noise is introduced in the *TB*s during the pattern recognition process.

[57] Figure 13 represents the covariance of the a priori errors for temperature (Figure 13a) and water vapor (Figure 13b). The matrix in Figure 13c is the correlation matrix between the a priori errors on temperature and water vapor. For temperature, diagonal values, representing the variance of a priori uncertainties, are between 5 and 10 K which means that the uncertainty is around 2 or 3 K. The diagonal element on the water vapor covariance matrix can reach 0.03 which represents a standard deviation for the uncertainty of about 0.17 (17%). The correlations between the temperature and water vapor a priori errors range between -0.2 and 0.2, this could seem insignificant but when integrated over few layers, the significance increase.

[58] Similarly to the Jacobians described in section 4.1, the covariance $S_f$ also depends on the situation: It is natural to have higher errors for atmospheric states that are more complex and therefore more difficult to retrieve. For example, an atmospheric profile with a lot of inversions in the vertical is more difficult to retrieve than a profile that is "smooth." It is clear also that a situation with a Jacobian that is of lower quality (lower magnitudes, which implies a lower signal-to-noise ratio) is also more difficult to retrieve. In this study, this $S_f$ variability is not taken into account, we are using a global $S_f$ , this is the case anyway for most of the retrieval techniques.

### 4.3. Retrieval Error Estimate

[59] Equation (3) is used here to estimate retrieval uncertainties using the Jacobians of section 4.1 and the a priori of section 4.2.

[60] First, the uncertainty estimate is performed independently for temperature and water vapor. This means that the matrix $A$ in equation (3) is just the Jacobian of temperature ($A = Jac_{temp}$) or just the Jacobian in water vapor ($A = Jac_{wv}$). This means that the "corruption" of the observations by the uncertainties on the other variable is not taken into account, but also that the link between temperature and water vapor is not exploited either. The retrieval uncertainties are estimated by the square root of the diagonal element in matrix $Q$ in Figure 14 for temperature (Figures 14a–14d) and water vapor (Figures 14e–14h). (The covariance matrix of uncertainties $Q$ is not diagonal of course, but the overall uncertainties are estimated by taking only the square root of the diagonal elements that represent the variance of the errors.) For temperature and water vapor, results are shown for MHS (Figures 14a and 14e), AMSU-A (Figures 14b and 14f), IASI (Figures 14c and 14g), and all of them (Figures 14d and 14h). In order to measure the impact of the Jacobian variability, the estimation is performed on a diverse set of 100 situations from the database of section 4.2. The grey lines represent the stack of these 100 situations, the black line the averaged uncertainty.

[61] It can be noted that the retrieval of temperature using MHS is highly variable, contrarily to AMSU-A retrievals. Furthermore, MHS is better at around 600 hPa, and AMSU-A is better for pressures lower than 400 hPa. It is however surprising that MHS gives relatively comparable retrieval statistics for temperature compared to AMSU-A, MHS is dedicated to water vapor and AMSU-A to temperature; but this is the direct result of the Jacobians in Figures 10 and 11. IASI is always better, this is not surprising since it has a lot of channels with relatively comparable Jacobian magnitudes, which automatically increases the signal-to-noise ratio. The "all" configuration that uses the three instruments together appears to be very close to the IASI results which would mean that there is no synergy among the three instruments. For water vapor the conclusions are similar except that MHS and AMSU-A retrievals are comparable, but again, less variable for AMSU-A. Again, no synergy seems to act. All these statistics can be compared to the a priori (section 4.2) that are roughly 2–3 K for temperature and $\simeq$17% for water vapor, depending on the atmospheric layer. This shows that the information on temperature has increased, but that for water vapor, only IASI instrument can

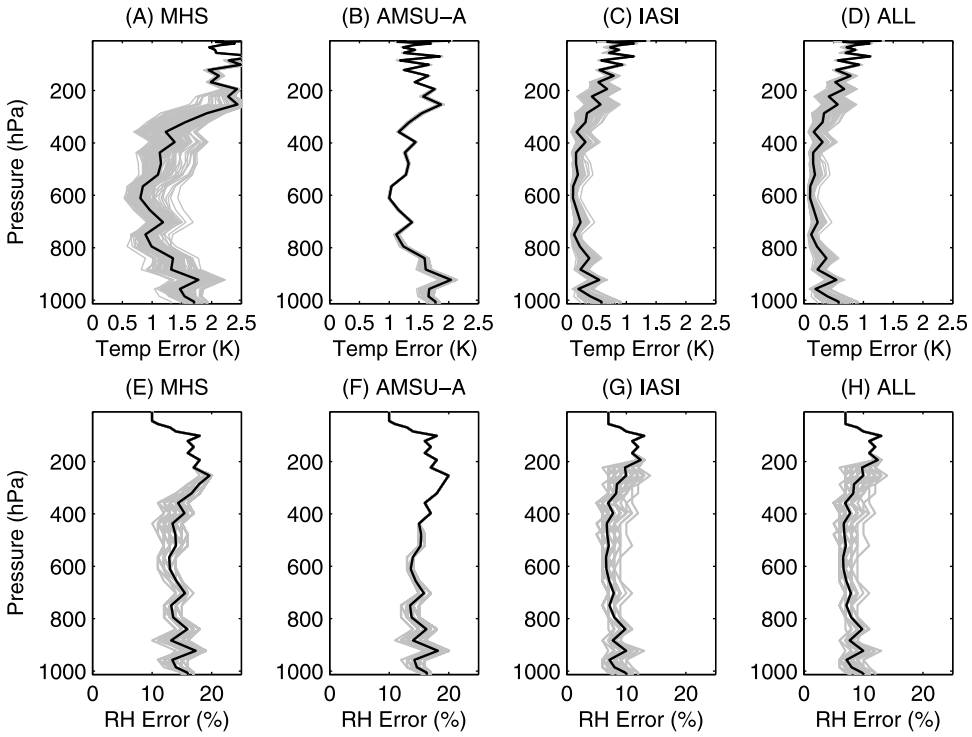


**Figure 14.** RMS error retrieval uncertainties for (a–d) temperature and (e–h) water vapor. For temperature and water vapor, results are shown for MHS (Figures 14a and 14e), AMSU-A (Figures 14b and 14f), IASI (Figures 14c and 14g), and all of them (Figures 14d and 14h). The grey lines represent these estimations for 100 situations, and the black line represents the averaged uncertainty. The estimation is performed independently for temperature and water vapor.

significantly improve the a priori information (this is really dependent on the vertical resolution that is chosen).

[62] It is important to note that these uncertainty estimates consider only one geophysical variable, temperature or water vapor. However, the Jacobians in section 4.1 show that channels can be sensitive to both of them. When estimating the uncertainties only for one geophysical variable at a time, the variability of the second one, that introduces additional uncertainties, is not taken into account. In order to measure the impact of combining temperature and water vapor for the estimation of retrieval uncertainty, similar experiments have been conducted on combined formulas. This means that the matrix $A$ in equation (3) includes the temperature and the water vapor Jacobians:

$$A = \left( \frac{\partial TB}{\partial T} \; \frac{\partial TB}{\partial WV} \right)$$

and that the a priori $S_f$ includes the covariance errors for temperature and water vapor together with their correlation (Figure 13):

$$S_f = \begin{pmatrix} cov(T,T) & cov(T,WV) \\ cov(WV,T) & cov(WV,WV) \end{pmatrix},$$

where $cov(T,T)$ is the covariance matrix of the FG errors (i.e., a priori matrix in Figure 13a), $cov(WV,WV)$ is in Figure 13b, and $cov(WV,T)$ is the covariance matrix of the FG errors on water vapor and temperature (that results in $cov(T,T)$, $cov(WV,WV)$ and the correlation matrix in Figure 13c). The statistics are different: First, the variability of the estimates seems to be lower in some configurations, in particular for the retrieval of water vapor with MHS instrument (not shown). The major impact is on the mean uncertainties themselves: Figure 15 represents the comparison of the uncertainty estimates for the individual estimate for temperature (Figure 15a), the combined estimate for temperature (Figure 15b), the individual estimate for water vapor (Figure 15c), and the combined estimate for water vapor (Figure 15d). For each configuration, the results are provided when using only MHS instrument, only AMSU-A, only IASI, and using all of them. The results represented here are the averaged uncertainty estimates, similarly to what was done in Figure 14.

[63] There are two essential conclusions from these experiments: First, the impact of combining both temperature and water vapor for the estimation of retrieval uncertainties is very important. The uncertainties in the retrieval of temperature increase relatively significantly, for IASI instrument, by 0.1/0.3 K. The uncertainties for the retrieval of water vapor have changed also: The IASI retrieval is degraded again when temperature and water vapor are combined, and the retrieval from MHS or AMSU-A became very close. Since we have both temperature and water vapor Jacobians, it is possible to combine them, but in order to be exhaustive,

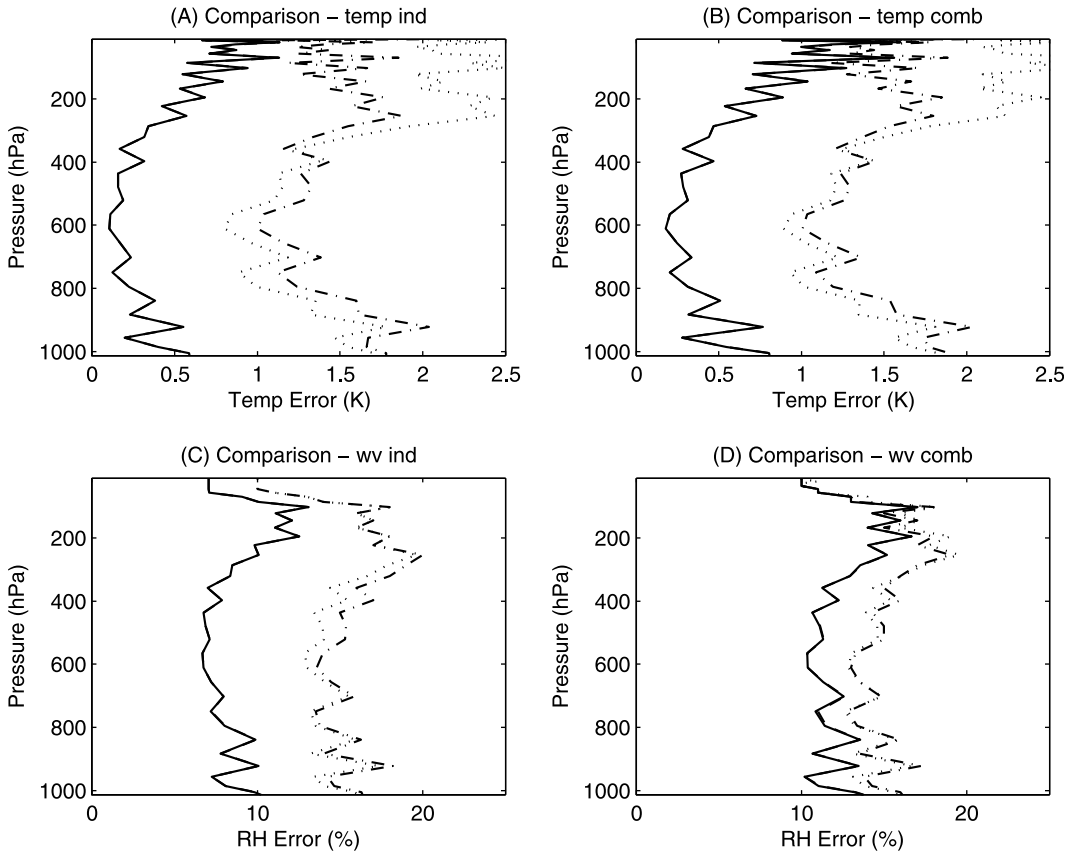


**Figure 15.** Comparison of the RMS error retrieval uncertainty estimates for (a) the individual estimate for temperature, (b) the combined estimate for temperature, (c) the individual estimate for water vapor, and (d) the combined estimate for water vapor. For each configuration, the results are provided when using only MHS instrument (dotted lines), only AMSU-A (dash-dotted lines), only IASI (dashed lines), and using all of them (solid lines). The IASI results are mostly obscured by the results using all measurements (solid line).

the Jacobians for all geophysical variable impacting the satellite measurements would be necessary (for example ozone or $N_2$ $O$ for IASI instrument).

[64] Since we know that the IASI channels are dependent on both temperature and water vapor, performing the retrieval of only the temperature is not correct. Instead of estimating the information content when both temperature and water vapor are retrieved simultaneously, another solution would be to estimate the additional noise in the measurements coming from the errors on the water vapor. This would be a more judicious comparison to the simultaneous retrieval. However, the uncertainties on the water vapor are not independent to the uncertainties on the temperature retrieval, this means that expression in equation (3) would not apply.

[65] The second striking result provided by these comparisons is that there appears to be no synergy: IASI is always better than the MW retrievals, and using all of them together does not improve the results. The magnitude of IASI or MW Jacobians is comparable, but IASI possesses so many channels that the signal-to-noise ratio dominates and cannot be impacted by the microwave observations.

### 4.4. Discussion on the Information Content Approach

[66] Two questions arise from the experiments of this section: Are the uncertainty estimates from the Information Content (IC) analysis realistic? Can this technique be used to measure synergy?

[67] It has been shown that many hypotheses need to be made to optimally use equation (3) of IC: Gaussian character of stochastic variables, linearization of the RTM around the first guess, independence of the FG with respect to observations, and unbiased RTM. Section 2 in this paper uses a synthetic application with a model that is linear with Gaussian statistics: Traditional IC is perfectly adapted to this case because it is based on the IC assumptions. For our more complex problem of the retrieval of temperature and water vapor from IR/MW instrument, the classical IC does not seem to provide realistic estimates of the retrieval uncertainties: In the work by *Divakarla et al.* [2006] the temperature and water vapor retrieved with the AIRS (Atmospheric InfraRed Sounder) instrument, which has similar characteristics to IASI, has been validated against radiosonde measurements and forecasts. The results obtained are a higher than the IC theoretical estimates, around 1–1.5K for temperature and more than 20% for water vapor. It is expected that results would be more positive for cloudy situations where the microwave observations would be more informative and complementary with respect to infrared (this will be the subject of a forthcoming study). One of the reasons why the classical IC appears to be deficient in this experiment is related to the hypothesis on

the IASI channel noise. It is assumed here that the instrument noise is independent from one channel to another. This is a simplification but without any official specification on this, it is not possible to something else (some more discussion on this will be provided by *Aires et al.* [2011]).

[68] IC is also very sensitive to the way it is used: It has been shown that when a Jacobian is used, the sensitivity of the observations to other constituents is often not taken into account, which can have strong impacts. For example, considering temperature and water vapor at the same time degrades the estimates, which is not a normal behavior for a linear retrieval scheme. Even if all the required hypotheses were verified, we have seen that all the components in equation (3) are situation dependent: The Jacobians $A$, the observational noise $S_\varepsilon$ [*Aires et al.*, 2002] and the a priori error covariance. This is a normal behavior but it means that in order to obtain more realistic estimates, the technique must be used with much more care by performing statistics in a representative sample of restitutions (i.e., what we did here).

[69] The main advantage of IC is its simplicity; this method allows for the comparison of different retrieval assumptions, but it can hardly provide good absolute uncertainty estimates. Since the retrieval estimates are not reliable, it would be questionable to use it to measure synergy. For example, in the IR/MW application of this study, the IR estimates are too optimistic, and therefore, no synergy with the MW observations is observed.

[70] Taking into account all the mentioned factors, getting realistic estimates would considerably complicate the use of IC and the main argument for its use, i.e., its simplicity, would be lost. It seems to be more pertinent to measure the synergy directly using real retrieval algorithms. This will be the objective of the companion paper [*Aires et al.*, 2011].

## 5. Conclusion and Perspective

[71] Satellite platforms (ENVISAT, MetOp) are today equipped with an increasing number of instruments, in order to obtain observations of a same location from multiple sensors, at different wavelengths (from the UV to the microwave), with different geometries (nadir or limb), and possibly in different modes (active and passive). Successions of several platforms that follow each other very closely, such as the A-train, also aim at observing a given area with even more sensors. Despite this wealth of coincident (or quasi-coincident) observations of a same atmospheric state with different perspectives, there are still limited attempts to benefit from the potential synergy between the information. Using the synergy is often suggested but rarely actually exploited. Most of the time, retrievals are performed independently for each instrument and the retrieved products are merged a posteriori. In this context, there is a strong need to study the theory of synergetic processes.

[72] This paper tried to explain some synergy mechanisms, how they occur and how to use them. The first conclusions are as follows:

[73] 1. Classical information content analysis can hardly be used to measure synergy.

[74] 2. There exist different types of synergy (additive, unmixing, indirect or denoising); each one can have a strong impact on the retrievals. Direct additive synergy that exploits the central limit theorem is not the same as unmixing synergy which is based on the introduction of a constrain in the inverse problem or the indirect synergy that can result sometimes in counterintuitive behaviors. It has been shown that additive synergy was important for our application, and that the indirect one was acting only for water vapor retrieval, not temperature. Our taxonomy helps formalize the inverse problems, this facilitates the elaboration of the retrieval schemes and the communication around them. The experiments performed with simple linear model illustrated well how to use all the a priori information in the retrieval problem.

[75] The application focused on the retrieval of the atmospheric temperature and water vapor profiles using three instruments (AMSU-A and MHS for the microwave, IASI for the infrared domain) all on board the MetOp platform. Experiments were conducted over ocean and under clear sky conditions. These conditions are not optimal to study synergy: the contribution of MW observations would be more important for cloudy cases. However, it is important to start this type of study with the simpler case, and increase complexity in the future. We showed that information content is very sensitive to the hypotheses that are used and that, in order to obtain realistic uncertainty characterization, it is necessary to obtain truly state-dependent statistics. The simplicity of information content analysis is lost. Since this technique is not truly reliable to obtain absolute retrieval uncertainties, its use to measure the synergy when combining various satellite observations is not reliable. In our application, no synergy between infrared and microwave observations was found because of too optimistic IR uncertainties.

[76] *Aires et al.* [2011] will show that it is, overall, easier to use directly real statistical inversion schemes to measure synergy. Furthermore, infrared and microwave observations will show strong synergy to retrieve temperature and water vapor even under clear-sky conditions.

[77] **Acknowledgments.** We are grateful to Cyril Crevoisier and Raymond Armante for providing IASI simulations using the 4A model. We would like to express our gratitude for Catherine Prigent for interesting comments and discussions. This project has been funded by ESA's General Studies Programme (GSP) under contract 21837/08/NL/HE, "Towards a synergetic approach for the retrieval of atmospheric geophysical parameters from optical/infrared and microwave measurements."

## References

Aires, F., and C. Prigent (2007), Sampling techniques in high-dimensional spaces for the development of satellite remote sensing database, *J. Geophys. Res.*, *112*, D20301, doi:10.1029/2007JD008391.

Aires, F., C. Prigent, W. B. Rossow, and M. Rothstein (2001), A new neural network approach including first-guess for retrieval of atmospheric water vapour, cloud liquid water path, surface temperature and emissivities over land from satellite microwave observations, *J. Geophyhs. Res.*, *106*(D14), 14,887–14,907.

Aires, F., A. Chédin, N. Scott, and W. B. Rossow (2002), A regularized neural network approach for retrieval of atmospheric and surface temperatures with the IASI instrument, *J. Appl. Meteorol.*, *41*(2), 144–159.

Aires, F., M. Paul, C. Prigent, B. Rommen, and M. Bouvet (2011), Measure and exploitation of multisensor and multiwavelength synergy for remote sensing: 2. Application to the retrieval of atmospheric temperature and water vapour from MetOp, *J. Geophys. Res.*, doi:10.1029/2010JD014702, in press.

Chalon, G., F. Cayla, and D. Dievel (2001), IASI: an advanced sounder for operational meteorology, paper presented at the 52nd Congress, Int. Astronaut. Fed., Toulouse, France, 1–5 Oct.

Chéruy, F., and F. Aires (2009), Cluster analysis of cloud properties over the southern Europe Mediterranean area in observations and a model, *Mon. Weather Rev.*, *137*(10), 3161–3176.

Cho, C., and D. H. Staelin (2006), Cloud clearing of Atmospheric Infrared Sounder hyperspectral infrared radiances using stochastic methods, *J. Geophys. Res.*, *111*, D09S18, doi:10.1029/2005JD006013.

Deblonde, G., and S. J. English (2001), Evaluation of the FASTEM-2 fast microwave oceanic surface emissivity model, paper presented at the Eleventh International TOVS Study Conference, World Meteorol. Organ., Budapest, 20–26 Sept.

Divakarla, M. G., C. D. Barnet, M. D. Goldberg, L. M. McMillin, E. Maddy, W. Wolf, L. Zhou, and X. Liu (2006), Validation of Atmospheric Infrared Sounder temperature and water vapor retrievals with matched radiosonde measurements and forecasts, *J. Geophys. Res.*, *111*, D09S15, doi:10.1029/2005JD006116.

Eyre, J. R. (1991), A fast radiative transfer model for satellite sounding systems, *ECMWF Res. Dep. Tech. Memo.*, *176*, Eur. Cent. for Med.-Range Weather Forecasts, Reading, U. K.

Hewison, T. J., and R. Saunders (1996), Measurements of the AMSU-B antenna pattern, *IEEE Trans. Geosci. Remote Sens.*, *34*(2), 405–412.

Jakob, C., G. Tselioudis, and T. Hume (2005), The radiative, cloud, and thermodynamic properties of the major tropical western Pacific cloud regimes, *J. Clim.*, *8*, 1203–1215, doi:10.1175/JCLI3326.1.

Kalnay, E. (2003), *Atmospheric Modeling, Data Assimilation and Predictability*, 364 pp., Cambridge Univ. Press, Cambridge, U. K.

Karbou, F., C. Prigent, L. Eymard, and J. Pardo (2005), Microwave land emissivity calculations using AMSU measurements, *IEEE Trans. Geosci. Remote Sens.*, *43*, 948–959.

Li, J., W. P. Menzel, W. Zhang, F. Sun, T. J. Schmit, J. J. Gurka, and E. Weisz (2004), Synergistic use of MODIS and AIRS in a variational retrieval of cloud parameters, *J. Appl. Meteorol.*, *43*, 1619–1634.

Lloyd, S. (1992), Least squares quantization in PCM, *IEEE Trans. Inf. Theory*, *28*, 129–137.

Matricardi, M., F. Chevallier, G. A. Kelly, and J.-N. Thépaut (2004), An improved general fast radiative transfer model for the assimilation of radiance observations, *Q. J. R. Meteorol. Soc.*, *130*, 153–173.

Mo, T. (1996), Prelaunch calibration of the Advanced Microwave Sounding Unit-A for NOAA-K, *IEEE Trans. Microwave Theory Tech.*, *44*, 1460–1469.

Moody, J., and C. Darken (1989), Fast learning in networks of locally-tuned processing units, *Neural Comput.*, *1*, 281–294.

Rabier, F., A. McNally, E. Andersson, P. Courtier, P. Undein, J. Eyre, A. Hollingsworth, and F. Bouttier (1998), The ECMWF implementation of three-dimensional variational assimilation (3D-Var). II. Structure functions, *Q. J. R. Meteorol. Soc.*, *124*, 1809–1829.

Rodgers, C. D. (1990), Characterization and error analysis of profiles retrieved from remote sounding measurements, *J. Geophys. Res.*, *95*, 5587–5595.

Rodgers, C. D. (2000), *Inverse Methods for Atmospheric Sounding—Theory and Practice*, World Sci., London.

Saunders, R. W., M. Matricardi, and P. Brunel (1999), An improved fast radiative transfer model for assimilation of satellite radiance observations, *Q. J. R. Meteorol. Soc.*, *125*, 1407–1425.

Scott, N. A. (1974), A direct method of computation of transmission function of an inhomogeneous gaseous medium: Description of the method and influence of various factors, *J. Quant. Spectrosc. Radiat. Transfer*, *14*, 691–707.

Scott, N. A., and A. Chédin (1981), A fast line-by-line method for atmospheric absorption computations: The Automatized Atmospheric Absorption Atlas, *J. Appl. Meteorol.*, *20*, 556–564.

Simmons, A. J., and J. K. Gibson (2000), The ERA-40 project plan, *ERA-40 Proj. Rep. Ser.*, *1*, 63 pp., Eur. Cent. for Med.-Range Weather Forecasts, Reading, U. K.

Susskind, J., C. D. Barnet, and J. M. Blaisdell (2003), Retrieval of atmospheric and surface parameters from AIRS/AMSU/HSB data in the presence of clouds, *IEEE Trans. Geosci. Remote Sens.*, *41*(2), 390–409.

Tarantola, A. (1987), *Inverse Problem Theory. Models for Data Fitting and Model Parameter Estimation*, Elsevier, Amsterdam.

F. Aires, Estellus, 93 bl Sébastopol, F-75002 Paris, France. (filipe.aires@estellus.fr)